

## Nucleic Acid Ladders

### CROSS REFERENCE TO RELATED APPLICATION

5 The present application <sup>is</sup> ~~is~~ claims priority to U.S. Provisional Application No. 60/040,914, filed March 27, 1997, the contents of which are entirely incorporated herein by reference.

### FIELD OF THE INVENTION

10 The present invention is in the field of molecular biology. In particular, the invention relates to determining the size of nucleic acid molecules by comparing such nucleic acid molecules to a nucleic acid marker ladder (single- or double-stranded) having a known size.

### BACKGROUND OF THE INVENTION

#### Electrophoresis

15 A common method of analyzing nucleic acid (*e.g.*, DNA or RNA) fragments is by separation in an agarose or polyacrylamide gel matrix. Such matrices provide a medium in which such fragments of different sizes and forms may be analyzed. In the presence of an electrical current, nucleic acid fragments will migrate through such a gel matrix in the direction of the positive electrode. Since small, linear nucleic acid molecules migrate more easily and quickly through the pores of the gel than do larger molecules, the matrix acts as a molecular sieve, separating fragments of different sizes. Larger nucleic acid fragments are typically electrophoresed on low concentration agarose gels, while smaller fragments are separated on higher concentration agarose gels or on polyacrylamide gels, since polyacrylamide has a higher resolution capacity than agarose.

20

### Detection of Nucleic Acid Fragments

Detection of nucleic acid molecules electrophoresed in agarose or polyacrylamide gels can be accomplished by a variety of techniques, including the use of fluorescent dyes such as ethidium bromide and SYBR Green. These dyes bind to the nucleic acid molecules and fluoresce when exposed to UV light. Alternatively, the nucleic acid molecules can be detectably labeled by chemically coupling them with radioactive, fluorescent or chemiluminescent labels.

### Sizing Standards

For the purpose of analyzing nucleic acid fragments on agarose or polyacrylamide gels, nucleic acid molecular weight standards are very useful tools, providing the researcher with a point of reference for estimating the quality, size, and/or quantity of the nucleic acid sample. A standard is typically run simultaneously on the gel with the sample (*e.g.*, in parallel with the sample), and following detection, a comparison is made between the sample band(s) and the bands of the standard. Knowing the size (in base pairs) of the standard allows the size of the unknown fragment(s) to be estimated.

Common standards used for estimating the size of nucleic acid fragments on gels include naturally occurring genomic DNA of bacteriophages (*e.g.*, lambda bacteriophage and  $\phi$ X174). To prepare such standards, the phage genomic DNA is cleaved into a population of nucleic acid fragments of known size using a specific restriction endonuclease. These types of standards are commonly called "nucleic acid markers" (*e.g.*, "DNA markers").

Another type of nucleic acid standard is produced by engineering plasmids to contain recognition/cleavage sites for one or more specific restriction endonucleases at particular intervals in the plasmid. *See, e.g.*, WO 95/11971. Upon digestion of the plasmid with the specific endonuclease(s), nucleic acid fragments of specific known sizes are generated. For accuracy in size determination and for ease of use, it is beneficial to have numerous bands that

increase in size in regular, even intervals. The bands of the standard should also be able to be detected with equivalent intensity. These types of standards are commonly called "nucleic acid ladders" (e.g., "DNA ladders").

5 A 10 bp DNA ladder which consists of a set of twenty fragments containing repeats of a 10 bp sequence is commercially available from Life Technologies, Inc. (Gaithersburg, Maryland). This ladder is constructed by ligation of synthetic double-stranded oligonucleotides, and is not generated from a plasmid. The multimers in the ladder contain multiple deoxyuracil bases in one strand and can be converted to a single-stranded form by incubation with uracil DNA glycosylase (UDG), which degrades the uracil-containing strand. Bands at 10 bp and 100 bp are highlighted for orientation within the standard.

10 This 10 bp DNA ladder can be used for sizing double-stranded DNA, or treated with UDG as described above for sizing single-stranded DNA, and may be stained with ethidium bromide or end-labeled with T4 polynucleotide kinase. However, due to the nature of the production procedure (ligation of oligonucleotides), it is not amenable to fluorescent detection, nor is it useful for Amplification Fragment Length Polymorphism (AFLP) analysis (Vos *et al.*, *Nucl. Acids. Res.* 23:4407-4414 (1995)), primarily because the resulting range of fragment sizes is too low and the electrophoretic bands are not sharply resolved.

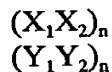
15 Other commercially available products include a 20 bp DNA ladder sold by GenSura, a 20 bp DNA ladder sold by Invitrogen (these may be the same product sold by the two companies), and LTI's 25 bp ladder. The LTI 25 bp DNA ladder is derived from a plasmid engineered to contain several repeats of a 25 bp sequence which can be digested by restriction enzymes to generate bands increasing in size by 25 bp. The repeated sequence in the LTI 25 bp DNA ladder, however, can be used only for sizing double-stranded DNA.

20

25

## BRIEF SUMMARY OF THE INVENTION

The present invention relates to a nucleic acid molecule comprising two or more copies of a repeat-containing sequence having the formula



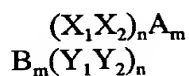
wherein:

$X_1$  and  $X_2$  are single nucleotides or derivatives thereof which may or may not be identical;

$Y_1$  and  $Y_2$  are single nucleotides or derivatives thereof complementary to  $X_1$  and  $X_2$ , respectively; and

$n$  is an integer from 1 to  $1 \times 10^{10}$ ,

and wherein two or more of the repeat-containing sequence copies are separated by a restriction site and the nucleotide compositions of  $(X_1X_2)_n$  and  $(Y_1Y_2)_n$  are substantially identical. The invention also relates to a nucleic acid molecule comprising two or more copies of a repeat-containing sequence having the formula



wherein:

$X_1$  and  $X_2$  are single nucleotides or derivatives thereof which may or may not be identical;

$Y_1$  and  $Y_2$  are single nucleotides or derivatives thereof complementary to  $X_1$  and  $X_2$ , respectively;

$n$  is an integer from 1 to  $1 \times 10^{10}$ ;

$A$  is a nucleotide or a derivative thereof;

$B$  is a nucleotide or a derivative thereof; and

$m$  is an integer from 1 to 100,

wherein two or more of the repeat-containing sequence copies are separated by a restriction site and wherein the nucleotide compositions of  $(X_1X_2)_nA_m$  and

$B_m(Y_1Y_2)_n$  are substantially identical. The invention is particularly directed to such nucleic acid molecules which are DNA molecules, RNA molecules, or DNA/RNA hybrid molecules, and to such molecules which may be double-stranded.

In another aspect, the invention is directed to such molecules wherein two or more of the repeat-containing sequence copies are separated by cleavage at the restriction site. In preferred such aspects, 2 to 500, 2 to 100 or 2 to 50 of said repeat-containing sequence copies are separated by cleavage at said restriction site. The invention is also directed to such separated repeat-containing sequence copies which are detectably labeled or which are subjected to conditions favoring the conversion of the sequences into a single-stranded form which may themselves be detectably labeled. Preferred detectable labels for this aspect of the invention include radiolabels, fluorescent labels and chemiluminescent labels.

The invention is also directed to the above-described nucleic acid molecules, wherein each of the repeat-containing sequence copies are the same length. In particular, the invention is directed to such nucleic acid molecules wherein each of the repeat-containing sequence copies has a length of from about 5 nucleotides to about 1000 nucleotides and more preferably a length of from about 5 nucleotides to about 100 nucleotides.

The invention is also directed to the above-described nucleic acid molecules wherein  $(X_1X_2)$  and  $(Y_1Y_2)$  are identical, and to those wherein the repeat-containing sequence is a palindromic sequence.

The invention is also directed to the above-described nucleic acid molecules wherein the nucleotides or derivatives thereof are selected from the group consisting of dUTP, dATP, dTTP, dCTP, dGTP, dITP, 7-deaza-dGTP,  $\alpha$ dATP,  $\alpha$ dTTP,  $\alpha$ dGTP,  $\alpha$ dCTP, ddATP, ddTTP, ddCTP and ddGTP. In additional preferred nucleic acid molecules of the invention, n is an integer from 1 to  $1 \times 10^6$ . In other preferred nucleic acid molecules of the invention, n is an integer from 1 to  $1 \times 10^3$  and m is an integer from 1 to 10.

The invention is also directed to vectors comprising the above-described nucleic acid molecules, particularly the plasmid vector pAH102.4. The invention is also directed to a host cell comprising the above-described nucleic acid molecules or vectors.

5           The invention is also directed to nucleic acid ladders comprising the above-described nucleic acid molecules. The invention also provides methods for making additional nucleic acid ladders, a preferred such method comprising

(a)       mixing the above-described nucleic acid molecule with a restriction enzyme which cleaves at the restriction site; and

10           (b)       incubating the mixture under conditions favoring the cleavage of the nucleic acid molecule at one or more of the restriction sites. Preferably, such cleavage is a partial digestion to form a population of multimers (*e.g.*, monomer repeat, dimer repeat, trimer repeat, tetramer repeat, etc.), thereby forming a ladder. In a preferred such aspect, the invention provides such methods further comprising treating the nucleic acid molecule under conditions favoring the  
15           conversion of the nucleic acid molecule into a single-stranded form.

The invention also provides methods for determining the size of a nucleic acid molecule, a preferred such method comprising

20           (a)       separating the nucleic acid ladder of the invention, and the nucleic acid molecule to be sized, according to size; and

(b)       determining the size of the nucleic acid molecule by comparison to the nucleic acid ladder.

25           The invention also relates to kits comprising a carrier means, such as a box, carton or the like, being compartmentalized to receive in close confinement therein one or more container means, such as tubes, vials, ampules, bottles or the like, wherein a first container means comprises the nucleic acid molecule or ladder of the present invention.

## BRIEF DESCRIPTION OF THE DRAWINGS

**Figure 1** depicts a plasmid map of pAH102.4 described in Example 1.

**Figure 2** depicts a gel showing a comparison of the 10 bp ladder of the present invention to various commercially available ladders. Lane 1: 25 bp commercial ladder (Life Technologies, Inc.). Lane 2: 10 bp commercial ladder (Life Technologies, Inc.). Lane 3: 10 bp ladder of the present invention. Lane 4: 20 bp Superladder-Low™ commercial ladder (GenSura). Lane 5: 20 bp commercial ladder (Invitrogen).

**Figure 3** depicts a <sup>32</sup>P-labeled single-stranded 10 bp ladder of the invention (duplicate samples in lanes 1 and 2).

## DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to nucleic acid molecules which may be used as standards for estimating the size (in base pairs) and mass of linear, double-stranded or single-stranded nucleic acid molecules separated by size, preferably by electrophoresis on agarose or polyacrylamide gels. The nucleic acid molecules of the invention may be DNA molecules, RNA molecules or DNA/RNA hybrid molecules, and may be double-stranded or single-stranded.

In particularly preferred aspects of the invention, the nucleic acid molecules are produced from vectors designed to contain at least two copies, preferably about 2 to about 500 copies, 20 to 100 copies or 2 to 50 copies, and most preferably about 200 to about 300 copies, of a repeat-containing sequence. The repeat-containing sequences contained in the nucleic acid molecules of the invention are preferably about 5 to about 1000 nucleotides (for single-stranded nucleic acid molecules) or base pairs (for double-stranded nucleic acid molecules) in length, and most preferably about 5-100 nucleotides or base pairs in length. In a preferred aspect of the invention, each of the repeat-containing sequences is of

the same size (*i.e.*, length). In another preferred aspect of the invention, the repeat-containing sequences have the same base compositions in the top and bottom strands. The repeat-containing sequences having the same base composition may be palindromic sequences, where the sequence of the top and bottom strands are mirror images. According to the invention, the top and bottom strands of the repeat-containing sequence need only be substantially identical in base composition. By "substantially identical in base composition" is meant that the top and bottom strands of the repeat-containing sequence are about 80%, preferably at least about 90%, more preferably at least about 95%, still more preferably at least about 98% or about 99%, and most preferably at least about 100%, identical in base composition. Although substantially identical base compositions include palindromic sequences, the order of the sequence in the top and bottom strand need not be the same according to the invention.

Example repeat-containing sequences of the invention include but are not limited to:

ATCTCAGGAT (SEQ ID NO:1)  
TAGAGTCCTA

ATCAGTCGAT (SEQ ID NO:2) 4/10 ID  
TAGTCAGCTA

ATCGCATGAT (SEQ ID NO:3)  
TAGCGTACTA

ATCATGCGAT (SEQ ID NO:4) 4  
TAGTACGCTA

According to this aspect of the invention, a particularly preferred nucleic acid molecule contains about 231 repeats, each of which is about 5-100 nucleotides or base pairs in length. The size of the repeats is preferably selected to be a multiple of an integer, wherein the integer is two or more. In one preferred



embodiment, the integer is 10, 50 or 100. More preferably, the repeat sequence is a 10 bp sequence.

In order to prepare the repeat-containing nucleic acid molecules of the invention, one may first prepare an oligonucleotide that contains multiple copies of a repeat. The oligonucleotide may be prepared by methods of solid phase synthesis or other methods suitable for synthesis of oligonucleotide molecules that will be apparent to one of ordinary skill in the art.

According to one aspect of the invention, a preferred such oligonucleotide may be represented by the Formula (I):



wherein:

$X_1$  and  $X_2$  are single nucleotides or derivatives thereof which may or may not be identical;

$Y_1$  and  $Y_2$  are single nucleotides or derivatives thereof complementary to  $X_1$  and  $X_2$ , respectively; and

$n$  is an integer from 1 to  $1 \times 10^{10}$ .

In particularly preferred nucleic acid molecules of this aspect of the invention,  $X_1$  and  $X_2$  are not identical. In other particularly preferred nucleic acid molecules of this aspect of the invention, two or more of the repeat-containing sequence copies represented by  $X_1X_2$  and  $Y_1Y_2$  are separated by a restriction site, and the nucleotide compositions of  $(X_1X_2)_n$  and  $(Y_1Y_2)_n$  are substantially identical. Preferably, the nucleotide compositions of  $(X_1X_2)$  and  $(Y_1Y_2)$  are the same. In other particularly preferred nucleic acid molecules of this aspect of the invention,  $n$  is an integer from 1 to  $1 \times 10^6$ .

According to another aspect of the invention, an alternative preferred oligonucleotide may be represented by the Formula (II):



wherein:

$X_1$  and  $X_2$  are single nucleotides or derivatives thereof which may or may not be identical;

$Y_1$  and  $Y_2$  are single nucleotides or derivatives thereof complementary to  $X_1$  and  $X_2$ , respectively;

$n$  is an integer from 1 to  $1 \times 10^{10}$ ;

$A$  is a nucleotide or a derivative thereof;

$B$  is a nucleotide or a derivative thereof; and

$m$  is an integer from 1 to 100.

In particularly preferred nucleic acid molecules of this aspect of the invention,  $X_1$  and  $X_2$  are not identical. In other particularly preferred nucleic acid molecules of this aspect of the invention, two or more of the repeat-containing sequences represented by  $X_1X_2$  and  $Y_1Y_2$  are separated by a restriction site, and the nucleotide compositions of  $(X_1X_2)_nA_m$  and  $B_m(Y_1Y_2)_n$  are substantially identical. Preferably, the nucleotide composition of  $(X_1X_2)_nA_m$  and  $B_m(Y_1Y_2)_n$  are the same. In other particularly preferred nucleic acid molecules of this aspect of the invention,  $n$  is an integer from 1 to  $1 \times 10^3$  and  $m$  is an integer from 1 to 10.

According to the invention, nucleotides or derivatives thereof suitable for preparing the oligonucleotides and nucleic acid molecules of the invention include, but are not limited to, dUTP, dATP, dTTP, dCTP, dGTP, dITP, ATP, TTP, UTP, GTP, CTP, 7-deaza-dGTP,  $\alpha$ dATP,  $\alpha$ dTTP,  $\alpha$ dGTP,  $\alpha$ dCTP, ddATP, ddTTP, ddCTP and ddGTP. Other nucleotides (deoxy and dideoxy) and derivatives thereof suitable for use in forming the nucleic acid molecules of the invention will be familiar to one of ordinary skill in the art.

The repeat-containing sequences of the invention may be ligated to produce multiple repeats. Since such repeats are separated by restriction sites according to the invention, one or more of these repeats may subsequently be separated by cleavage with a restriction enzyme such as a blunt-end or sticky-end restriction endonuclease. A restriction endonuclease is an enzyme that has the

capacity to recognize a specific base sequence (usually 4, 5, or 6 base pairs in length) in a DNA molecule, and to cleave the DNA molecule where this sequence appears. In the practice of the present invention, restriction enzymes and restriction sites should be chosen which give identical blunt-end or sticky-end fragments. Examples of blunt-end restriction enzymes suitable for use in the invention, and their cleavage sites, include without limitation:

5	<i>AluI</i>	5'-AG↓CT-3' 3-TC↓GA-5'
10	<i>DraI</i>	5'-TTT↓AAA-3' 3'-AAA↓TTT-5'
	<i>Eco47 III</i>	5'-AGC↓GCT-3' 3'-TCG↓CGA-5'
	<i>EcoRV</i>	5'-GAT↓ATC-3' 3'-CTA↓TAG-5'
15	<i>FspI</i>	5'-TGC↓GCA-3' 3'-ACG↓CGT-5'
	<i>HpaI</i>	5'-GTT↓AAC-3' 3'-CAA↓TTG-5'
20	<i>MscI</i>	5'-TGG↓CCA-3' 3'-ACC↓GGT-5'
	<i>NruI</i>	5'-TCG↓CGA-3' 3'-AGC↓GCT-5'
	<i>PvuII</i>	5'-CAG↓CTG-3' 3'-GTC↓GAC-5'
25	<i>RsaI</i>	5'-GT↓AC-3' 3'-CA↓TG-5'
	<i>ScaI</i>	5'-AGT↓ACT-3' 3'-TCA↓TGA-5'

*SmaI* 5'-CCC↓GGG-3'  
3'-GGG↓CCC-5'

*SspI* 5'-AAT↓ATT-3'  
3'-TTA↓TAA-5'

5 *StuI* 5'-AGG↓CCT-3'  
3'-TCC↓GGA-5'

*ThaI* 5'-CG↓CG-3'  
3'-GC↓GC-5'

10 *DraI* 5'-TTT↓AAA-3'  
3'-AAA↓TTT-5'

Examples of sticky-end restriction enzymes suitable for use in the invention, and their cleavage sites, include without limitation:

*AvaI* 5'-C↓PyCGPuG-3'  
3'-GPuGCPy↓C-5'

15 *BamHI* 5'-G↓GATCC-3'  
3'-CCTAG↓G-5'

*BanII* 5'-GPuGCPy↓C-3'  
3'-C↓PYCGPuG-5'

20 *BglII* 5'-A↓GATCT-3'  
3'-TCTAG↓A-5'

*ClaI* 5'-AT↓CGAT-3'  
3'-TAGC↓TA-5'

*EcoRI* 5'-G↓AATTC-3'  
3'-CTTAA↓G-5'

25 *HindIII* 5'-A↓AGCTT-3'  
3'-TTCGA↓A-5'

*HpaII* 5'-C↓CGG-3'  
3'-GGC↓C-5'

	<i>KpnI</i>	5'-GGTAC↓C-3' 3'-C↓CATGG-5'
	<i>MseI</i>	5'-T↓TAA-3' 3'-AAT↓T-5'
5	<i>NcoI</i>	5'-C↓CATGG-3' 3'-GGTAC↓C-5'
	<i>NdeI</i>	5'-CA↓TATG-3' 3'-GTAT↓AC-5'
10	<i>NotI</i>	5'-GC↓GGCCGC-3' 3'-CGCCGG↓CG-5'
	<i>PstI</i>	5'-CTGCA↓G-3' 3'-G↓ACGTC-5'
	<i>PvuI</i>	5'-CGAT↓CG-3' 3'-GC↓TAGC-5'
15	<i>SacI/SstI</i>	5'-GAGCT↓C-3' 3'-C↓TCGAG-5'
	<i>SalI</i>	5'-G↓TCGAC-3' 3'-CAGCT↓G-5'
20	<i>XbaI</i>	5'-T↓CTAGA-3' 3'-AGATC↓T-5'
	<i>XhoI</i>	5'-C↓TCGAG-3' 3'-GAGCT↓C-5'

The above-mentioned restriction enzymes, and others that may be equivalently used in the present invention, are available commercially, for example from Life Technologies, Inc. (Rockville, MD). See also Roberts, R.J., *Nucl. Acids Res.* 17(Suppl.):r347-r387 (1989), for other examples of restriction enzymes and their cleavage sites.

As will be understood by those of ordinary skill in the art, the nucleic acid molecules used to form the nucleic acid sizing ladder of the invention are

preferably linear or circular DNA molecules which are cleavable by a restriction enzyme. For example, the nucleic acid molecules may be derived from a chromosome, a vector, a cosmid, a plasmid or a viral genome. Preferably, the nucleic acid molecules are vector or viral molecules and derivatives thereof. The nucleic acids present in the vector or viral molecule may include exogenous nucleic acids which have been joined to produce the vector or viral molecule. In one preferred embodiment, the nucleic acid is DNA.

Preferably, the nucleic acid molecule contains an origin of replication (for example, *ori*) such that the nucleic acid molecule may autonomously replicate within a host cell. It is also preferable that the nucleic acid molecule contains a selectable or screenable marker. The origin of replication and the marker may be present on the same fragment. Host cells containing the nucleic acid molecule of the invention may be cultured and selected with a selection agent corresponding to the selectable marker.

According to the invention, the repeat sequences may be ligated into a vector which is then transformed into a host cell, the host cell is cultured, lysed and the vector isolated by well known techniques. The vector may then be cut with the restriction endonuclease that recognizes a restriction endonuclease site, thereby separating the repeats to give multiple copies of the repeat fragment. These multiple copies may be ligated together to form multimers, *e.g.*, dimers, trimers, tetramers, pentamers, hexamers and the like. Alternatively, partial digestion of the vector provides for multimers of varying sizes. Such multimers of variable size (*e.g.*, monomers, dimers, trimers, etc.) are used as sizing ladders according to the invention. Such sizing ladders may be double-stranded or single-stranded. Single-stranded ladders may be formed from double-stranded nucleic acid molecules or ladders of the invention by techniques that are well-known to one of ordinary skill in the art, such as heat or chemical denaturation.

As described in more detail in Example 1 below, preferred nucleic acid ladder may be produced from the plasmid pAH102.4 (Figure 1) by first digesting

the plasmid to completion with *StuI* to generate two copies of a 100 bp sequence and seven "cassettes" each containing 33 10-bp repeats (see Figure 1). The cassettes are then partially digested with *EcoRV* restriction enzyme to generate fragments increasing in size by increments of 10 bp, the largest fragment being 330 bp. The 100 bp fragment acts as an internal orientation marker within the ladder by increasing the brightness of the 100 bp band two to three times compared with the other bands in the ladder. Of course, the steps of complete digestion and partial digestion may be reversed in sequence.

The sizing ladder of the present invention can be used for estimating the size (in bp) of double-stranded nucleic acid (*e.g.*, DNA or RNA) fragments, preferably by electrophoresis on agarose or native polyacrylamide gels. This ladder can also be used to size single-stranded nucleic acid fragments when the strands are separated by heat or chemical denaturation. In particular, the 10 bp ladder containing the repeat sequence of SEQ ID NO:1 is useful as a standard for sizing linear, double-stranded and single-stranded nucleic acid fragments in the 10 bp to 330 bp range. The ladder of the present invention can also be detectably labeled, for example with a radiolabel, fluorescent label or chemiluminescent label as described below, and used as a standard in nucleic acid polymorphism analyses such as AFLP (EP 0 534 858; Vos, P. *et al.*, *Nucl. Acids Res.* 23(21):4407-4414 (1995)), Random Amplified Polymorphic DNA (RAPD; Williams, J.G.K. *et al.*, *Nucl. Acids Res.* 18(22):6531-6535 (1990)), DNA Amplification Fingerprinting (DAF; Caetano-Anollés *et al.*, *Bio/Technology* 9:553-557 (1991)) and microsatellite PCR or Directed Amplification of Minisatellite-region DNA (DAMD; Heath, D.D. *et al.*, *Nucl. Acids Res.* 21(24):5782-5785 (1993)). Alternatively, the ladder of the invention may be used as a reference in unlabeled versions of the above-described techniques.

The nucleic acid marker ladder of the present invention may be single-stranded or double-stranded. The double-stranded ladder is obtained directly from the double-stranded nucleic acid construct of the invention. Single strands may

be obtained by heating the double-stranded nucleic acid construct of the invention, or by treating it with a chaotropic agent or with a helicase. Alternatively, single strands will be obtained when separating the ladder on SDS-PAGE.

5 The migration of nucleic acid fragments on polyacrylamide gels is influenced by nucleotide base composition as well as size; that is, two nucleic acid bands containing the same number of base pairs but different nucleotide base compositions may migrate differently on polyacrylamide gels. According to the present invention, the base composition of the top strand of the repeat of the ladder is substantially the same as the base composition of the bottom strand, so  
10 the two nucleic acid strands will migrate identically when separated on denaturing polyacrylamide gels. Preferably, the base composition of the top and bottom strands are the same, for example, as with palindromic nucleic acid sequences. This is a major advantage over the known ladders which give rise to substantially different top and bottom single strands when separated.

15 The ladder bands can be detectably labeled by staining with ethidium bromide or SYBR Green, or by end-labeling using T4 polynucleotide kinase. Thus, another aspect of the invention relates to the ladder of the present invention which is detectably labeled with a stain or other detectable label. Labels suitable for detectably labeling the ladder of the invention include, but are not limited to,  
20 radiolabels (*e.g.*,  $^{32}\text{P}$ ,  $^{14}\text{C}$ ,  $^3\text{H}$  and the like), fluorescent labels (*e.g.*, fluorescein, rhodamine, phycocyanin and the like) and chemiluminescent labels (*e.g.*, using the PHOTO-GENE™ or ACES™ chemiluminescence systems, available commercially from Life Technologies, Inc., Rockville, Maryland).

25 In another embodiment, the present invention relates to a nucleic acid marker kit comprising a carrier means such as a box or carton having in close confinement therein at least one container means such as vials, tubes, jars, ampules and the like. A first container means may comprise the nucleic acid molecule or ladder of the present invention, in optionally labeled form. The first container means may also comprise a storage buffer such as about 10 mM TRIS-HCl (pH



8.0), about 1 mM EDTA and, optionally, about 50 mM NaCl. The nucleic acid marker ladder may be present at a concentration of about 1  $\mu$ M, and is preferably stored at about -20°C until use. A further container means may contain a reagent capable of detectably labeling the ladder of the present invention, such as ethidium bromide, SYBR Green, or T4 polynucleotide kinase.

Having now generally described the invention, the same will be more readily understood through reference to the following Examples which are provided by way of illustration, and are not intended to be limiting of the present invention, unless specified.

### Examples

#### *Example 1: Preparation of the Nucleic Acid Ladder*

The plasmid pAH102.4 (Figure 1) was prepared as follows. A 65 bp oligonucleotide molecule was synthesized having the following sequence (SEQ ID NO:5):

5'-AATTCTCGGG ATATCTCAGG ATATCTCAGG ATATCTCAGG  
3'-GAGCCC TATAGAGTCC TATAGAGTCC TATAGAGTCC

ATATCTCAGG ATATCTCGGG CTGCA-3'  
TATAGAGTCC TATAGAGCCC G-5'

The synthesized oligonucleotide was ligated into the pUC18 plasmid (Life Technologies, Inc., Rockville, MD) between the *Eco*RI and *Pst*I cloning sites to give plasmid pUC1810a. This plasmid was then cut with *Ava*I to generate a 50 bp DNA fragment containing five repeats of the 10 bp sequence (SEQ ID NO:1):

5'-ATCTCAGGAT-3'  
3'-TAGAGTCCTA-5'

A new plasmid, pAH102.1, was generated from pUC19 by retaining the origin of replication and *amp* genes of pUC19, and modifying its multiple cloning sites to contain restriction sites for *KpnI*, *StuI*, *EcoRV*, *AvaI*, *BanII*, *HindIII*, *EcoRI* and *BglII*. A series of repeats of the 50 bp fragment was inserted into the *AvaI* site of plasmid pAH102.1 by adding an excess amount of the 50 bp fragment in the ligation reaction mixture. A plasmid containing a hexamer of the 50 bp DNA was selected by direct size measurement of the DNA fragments of the resulting clones. This hexamer-containing plasmid was designated pAH102.2, and was cleaved with *BanII* to generate a 330 bp fragment. A series of repeats of this 330 bp fragment was inserted into the *BanII* site of pAH102.2 and a plasmid containing a heptamer of the 330 bp fragment and 42 copies of the 50 bp fragment was selected and designated as plasmid pAH102.3.

A DNA fragment containing three *StuI* restriction sites at 100 bp intervals was generated by amplifying phage  $\lambda$  DNA (Life Technologies, Inc., Rockville, Maryland) from position 31,384 to position 31,571 using the following primers:

Primer I: 5'-CAA CAA CAA GAA TTC AGG CCT AGC CAG TGC CTC GTC CAT TTT T-3' (SEQ ID NO:6)

Primer II: 5'-CAA CAA CAA AAG CTT CTT ACA TGG CCC AGG TGC AGT A-3' (SEQ ID NO:7)

This PCR-amplified  $\lambda$  DNA fragment was cleaved with *EcoRI* and *HindIII* and inserted into plasmid pAH102.3 between the engineered *EcoRI* and *HindIII* sites to give the desired plasmid, designated pAH102.4 (see Figure 1). The recombinant host cell comprising pAH102.4, *E. coli* STBL2(pAH102.4), was deposited on March 26, 1997, with the Collection, Agricultural Research Culture Collection (NRRL), 1815 North University Street, Peoria, Illinois 61604 USA, as Deposit No. NRRL B-21674.

The 10 bp ladder was produced from pAH102.4 by first digesting the plasmid to completion with *StuI*, thereby releasing two copies of a 100 bp

sequence and seven "cassettes" each containing 33 10 bp repeats. The cassettes were partially digested with *EcoRV*, carefully monitoring the enzyme concentration and incubation time such that the conditions favored the generation of fragments increasing in size in 10 bp increments, the largest fragment being 330 bp. In this ladder, the 100 bp fragment exhibited a density two to three times higher than that of the other bands in the ladder; this fragment thus acts as an internal orientation marker within the ladder.

***Example 2: Comparison of the 10 bp Ladder to Commercially Available Ladders***

To further characterize the nucleic acid ladder of the invention, the electrophoretic pattern of the double-stranded ladder was compared to that of several commercially available nucleic acid ladders. LTI's 25 bp ladder (1 µg, lane 1); LTI's 10 bp ladder (2 µg, lane 2); the 10 bp ladder produced according to Example 1 (2 µg, lane 3); GenSura's Superladder-Low™ 20 bp ladder (1 µg, lane 4), and Invitrogen's 20 bp ladder (0.8 µg, lane 5) were applied to a 4% LMP agarose gel containing 0.5 µg/ml ethidium bromide. The ladders were electrophoresed for 3.5 h at 75 V in a GIBCO BRL HORIZON® 11.14 apparatus.

As shown in Figure 2, the 10 bp ladder of the invention (lane 3) produced enhanced resolution, in 10 bp increments, of nucleic acid size compared to other commercially available nucleic acid sizing ladders. The present ladder also produced a 100 bp band of about 2- to 3-fold higher intensity than other bands, thus providing an internal sizing reference marker in the ladder of the invention. These results indicate that the double-stranded ladder of the invention facilitates rapid analysis of the size of double-stranded nucleic acid molecules with about 10 bp resolution.

***Example 3: Production and Characterization of Single-stranded Ladder***

To determine the utility of the present ladder in estimating the size of single-stranded nucleic acid molecules, a single-stranded ladder was prepared as described above and end-labeled with  $^{32}\text{P}$  using T4 kinase. Samples of this ladder were then electrophoresed as described above.

As shown in Figure 3, the  $^{32}\text{P}$ -labeled single-stranded ladder produced clearly resolved bands in the gel in 10 nucleotide increments. The range of sizes easily resolvable with this ladder was from about 10 nucleotides to about 330 nucleotides. These results demonstrate that the single-stranded ladder of the invention facilitates rapid analysis of the size of single-stranded nucleic acid molecules with about 10 bp resolution.

Having now fully described the present invention in some detail by way of illustration and example for purposes of clarity of understanding, it will be obvious to one of ordinary skill in the art that the same can be performed by modifying or changing the invention within a wide and equivalent range of conditions, formulations and other parameters without affecting the scope of the invention or any specific embodiment thereof, and that such modifications or changes are intended to be encompassed within the scope of the appended claims.

All publications, patents and patent applications mentioned in this specification are indicative of the level of skill of those skilled in the art to which this invention pertains, and are herein incorporated by reference to the same extent as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated by reference.